

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

Robust Identification and Control

M.A. Dahleh and J.N. Tsitsiklis

Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology
Cambridge, MA 02139

August 21, 1998

Abstract

This final report summarizes the research contributions under AFOSR grant No. F49620-95-1-0219. The work covered two major research directions. The first is in the area of robust linear and nonlinear control. In the linear area, a complete computationally-based methodology was developed for designing controllers that can meet multiple performance objectives in both the time and frequency domain. The research culminated in a book on multi-objective control. In the nonlinear area, an alternative to gain-scheduling that requires scheduling in Lyapunov space has been proposed which gives rise to a computational tool for synthesizing controllers with guaranteed stability. In addition, the theory of Neuro-dynamic programming was developed to handle large-scale nonlinear optimal control problems. This research culminated in another book on the theory and applications of Neuro-Dynamic programming. The second research direction is in the area of system identification. In that field, a new paradigm was proposed that allows deriving simple low-complexity models from noisy data obtained from complex systems. Within this paradigm, it is shown how to bridge the gap between stochastic and deterministic descriptions of noise. These developments have been shown to play a major role in many application domains.

19981113 055

Contents

1	Linear Robust Control	3
1.1	Introduction	3
1.2	Summary of Past Research	3
2	Computational Methods for Robust Nonlinear Control	6
2.1	Introduction	6
2.2	Summary of Past Research	7
3	Neuro-Dynamic Programming	10
3.1	Introduction	10
3.2	Summary of Past Research	11
4	Identification of Complex Systems	12
4.1	Introduction	12
4.2	Summary of Past Research	14
5	Technology Transfer	16
6	Major Publications	19
7	Theses Supervised	20

1 Linear Robust Control

1.1 Introduction

Much progress has been made in the area of linear robust control, and yet there are still many important issues to be addressed. We have been interested in developing one aspect of this theory, namely, ℓ_1 robust control. The ℓ_1 problem arises as the general disturbance rejection problem for linear time-invariant plants under bounded persistent disturbances, when the objective is to minimize the peak value of the error. Much of the progress is reported in [29] and references therein. In particular, our research in that area concentrated on developing efficient computational methods for constructing suboptimal controllers and simultaneously providing information on the structure of the optimal controller (e.g., the order of the optimal controller). In addition, we extended these algorithms to provide solutions for ℓ_1 problems with additional time-domain and frequency-domain constraints that arise in many practical problems. Since most of these problems are infinite dimensional, a solution usually consists of the following: deriving approximate methods with converging upper and lower bounds for the cost, providing methods for constructing suboptimal solutions, and providing structural information on the optimal controller [32, 33, 29, 25, 34, 35, 36, 147, 148, 37]. Parallel to our work, geometric methods for computing suboptimal ℓ_1 controllers were derived in [6] using dynamic programming. In addition, a state-space theory for ℓ_1 has emerged [109, 110, 12] exploiting viability theory. Although these approaches are seemingly different, we have recently shown that they can all be derived from dynamic programming arguments [38]. These results are still preliminary, and a complete theory with output feedback has not been derived. Also, the computational advantages of such approaches have not been investigated.

On a different end, tools for designing controllers that meet robust performance objectives have been entirely devoted to the case where performance is measured in terms of worst-case disturbance rejection [30, 25]. In many applications, however, performance objectives are stated in terms of the response of the system to a finite number of given inputs. An example of this kind of specifications is the problem of robust overshoot for step input commands. Only preliminary tools for addressing such problems have been developed [37, 63].

1.2 Summary of Past Research

We summarize below our research accomplishments in the area of robust control.

1. Computation of ℓ_1 Optimal Solutions

The contributions in this regard are marked by the introduction of the Delay Augmentation Algorithm for solving nonsquare problems (e.g., problems with more regulated variables than actuators) [32, 33]. This algorithm is based on squaring the system by introducing fictitious delayed inputs and outputs. The problem is solved iteratively as the number of delays increase. At each iteration, a square ℓ_1 problem is solved (the solution of which is known exactly). The main features of this algorithm are that: (1) at each iteration it gives upper and lower bounds for the

optimal objective function which are convergent; (2) it provides information about the structure of the controller; (3) it does not cause order inflation (it is not based on FIR approximations); (4) it involves solving one linear program iteratively. In many cases, the exact solution for nonsquare problems is provided.

For implementation purposes, all computations are performed using matrix algebra, often exploiting the structure of Toeplitz matrices resulting from convolution operators. An example of that is the development of methods for computing directions of zeros with multiplicity using Toeplitz matrix manipulation, without ever computing the Smith-McMillan Decomposition.

2. Controller Design for Mixed Objectives

In most applications, the controller is designed to meet several specifications, some in the time domain and others in the frequency domain. Some of these specifications are with respect to a fixed input, as opposed to a class of bounded signals. These problems cannot be systematically solved by formulating an appropriate \mathcal{H}_2 , \mathcal{H}_∞ or ℓ_1 problem. It is possible through weight selection to indirectly address some mixed objectives using one of the above three methods, however, this procedure remains ad hoc. Consequently, there was a great need to extend the above approaches to directly incorporate additional constraints.

Introducing additional linear constraints on the ℓ_1 problem results in infinite dimensional linear programs. These problems can only be solved approximately. To derive accurate bounds on the objective function, it is essential that we derive a dual problem that has the same objective value (i.e., has no duality gap). We have shown [35] that there is no duality gap for a large class of problems formulated as constrained ℓ_1 problems. An important class of constraints are those that give rise to linear matrix inequalities (LMIs). We have shown [147] that norm minimization problems with LMI constraints have dual representations without gaps, under mild assumptions on the constraints. These results include mixed objectives such as ℓ_1/\mathcal{H}_2 , $\ell_1/\mathcal{H}_\infty$, as well as general norm objectives with fixed input constraints. By approximating both primal and dual problems, we can approximate the objective function arbitrarily closely. In addition, details about the structure of the optimal controllers can be derived from the dual problem [34, 35, 147].

As a consequence of the solution of the mixed ℓ_1/\mathcal{H}_2 problem, we have recently proposed a new algorithm for solving the standard ℓ_1 problem that is based on splitting the cost into two components: the first is the ℓ_1 norm of the first N-taps of the closed loop response and the second is the \mathcal{H}_2 norm of the tail of the response. It was shown that this problem is equivalent to a finite-dimensional convex optimization problem, which can be readily solved and immediately provides converging upper and lower bounds of the optimal cost. This procedure is particularly efficient since it does not require computing interpolation conditions (i.e., exact linear constraints for the closed loop map to be feasible). It resembles the well-known Q-design procedure [14] in that it optimizes the Q-parameter directly, however, it also provides converging lower bounds. Details have been reported in [36].

3. Robustness Analysis and Synthesis

This area is concerned with the development of a computational theory to study directly uncertain plants. The uncertainty is structured in nature, possibly time varying. In this regard, we have built on the results in [26, 64] to come up with simple conditions for ℓ_∞ robustness analysis in the presence of structured uncertainty [25]. The conditions are stated in terms of the spectral radius of a matrix constructed from computing the ℓ_1 norms of certain closed loop maps. We have also analyzed the case of time-invariant perturbations when ℓ_∞ stability is required, and we have shown that the natural conditions are in the frequency domain (coincide with the standard μ results).

Since the spectral radius of a positive matrix can be computed by minimizing a scaled ℓ_1 norm, synthesis for structured uncertainty problems involves iterations between solving an ℓ_1 problem and finding optimal scales for the uncertainty. We have analyzed this algorithm in detail, and have shown its limitations. We have also proposed an alternative algorithm based on sensitivity analysis of the linear programming solution of the ℓ_1 problem [133].

4. Writing two Books on Robust Control

The book titled: *Control of Uncertain Systems: A Linear Programming Approach* written by Dahleh and Diaz-Bobillo presents a unified treatment of the theory of robust control design with emphasis on computational methods. It can serve as a starting point for researchers in the field as well as a textbook for a graduate class in control. In our opinion, this is the only book available that gives a comprehensive treatment of \mathcal{H}_2 , \mathcal{H}_∞ and ℓ_1 methods integrated in a *robust performance* framework, with emphasis on computations. As a follow-up to this book, a research monograph titled: *Computational Methods for Multi-Objective Control* is under development by Elia and Dahleh and has been accepted for publication. This book shows how generalized linear programs address a very wide range of practical control problems.

5. Software Development

A major part of this research, that parallels our research in computation, has been the development of software, which is currently available on the Internet. Using this software, we have studied a variety of benchmark problems (e.g., the X29 Aircraft, a flexible beam, a high purity distillation column). The following are the main new features of the software.

1. Mixed-objective problems are solved using several approaches. These include Delay Augmentation, finitely-many-variables, finitely-many-equations, and variations of Q-design (see [29]).
2. Software is interactive and additional time and frequency domain constraints can be graphically incorporated.
3. Characterizing feasible subspaces of closed loop maps by zeros. The computations involve lower triangular block-Toeplitz matrices.
4. All necessary computations are state-space based.
5. Optimization involves solving linear and convex programs.

2 Computational Methods for Robust Nonlinear Control

2.1 Introduction

The development of linear robust control has been accompanied by a parallel development of nonlinear control theory. In most of the existing nonlinear theory, no uncertainty is considered, and results are developed for systems of the form

$$\begin{aligned}\dot{x} &= f(x) + g(x)u, \\ y &= h(x)\end{aligned}\tag{1}$$

In the last decade, the study of such systems has provided an extension of various aspects of the linear theory, from general system theory [55], to problems of stabilization under smooth feedback [3], and optimal control [131, 57].

Among the many approaches for nonlinear controller design (e.g., see the control handbook [73]), a popular method that has resulted from this theory is based on dynamic inversion, primarily applied to feedback linearizable systems. Linear control techniques can be applied after linearization. It is well known, however, that this method suffers from several limitations. Sensitivity to parameter changes (lack of robustness), bandwidth constraints, and saturation constraints are some examples of the difficulties faced by this approach. In the presence of such constraints, a controller cannot remove the nonlinear dynamics of the process, or it may not be advantageous to do so.

This leads to fundamental open issues in nonlinear control: on the one hand, the development of “truly nonlinear” control designs which would characterize regions attraction of nonlinear systems and devise control strategies to keep systems within these regions. Secondly, the incorporation of robustness with respect to unmodeled dynamics, which has remained largely UN addressed by this theory.

A general approach which has been proposed recently to accomplish this objective has been the use of storage functions of the Lyapunov type. Such functions appear in many results of the nonlinear theory, such as the analysis of stability and \mathcal{H}_∞ -type performance [131], stabilization [3, 114], nonlinear \mathcal{H}_∞ -synthesis [57, 58, 82]. In its more general sense, one would find a robust-control-lyapunov function (RCLF) for the desired closed loop behavior [43], and the control design would be based on certain guaranteed decay rates for this function. Dynamic Inversion can be viewed as a special case of this approach (e.g. complete inversion with dynamics replaced by linear ones corresponds to imposing a quadratic RCLF). Optimal control problems are also a special case where the RCLF is derived from the Hamilton-Jacobi equations. From this point of view, it may not be necessary to find an inverse of the whole process, but rather it may be possible, indeed desirable, to retain dynamics that work in our favor.

While these results are promising and contribute to understanding the structure of nonlinear systems, they as yet have limited impact on practical problems, since finding such storage functions (e.g. a RCLF) is a nontrivial problem. The main difficulty is that while the computation of Lyapunov/storage functions for linear systems is a

tractable problem of the Linear Matrix Inequality (LMI) type [15], the extension to a general nonlinear system, e.g. of the form (1), gives a partial differential equation or inequality (e.g., a Hamilton-Jacobi equation) which is not easy to solve. Numerical approximation for these types of problems involves gridding the state space and therefore becomes intractable for moderately sized problems.

2.2 Summary of Past Research

Our previous work addresses three major avenues in nonlinear analysis and design.

1. Scheduling in Lyapunov Space

The objective here is to provide an alternative to gain-scheduling, using Robust Control Lyapunov Functions (RCLF) [42]. This addresses a class of systems known as Quasi-Linear-Parameter-Varying (LPV) that we will discuss shortly.

The problem setup is the following: one is given a nonlinear system influenced by a control signal u and a disturbance signal w , where $u(t) \in \mathcal{U} \subseteq \mathbf{R}^m$ and $w(t) \in \mathcal{W} \subseteq \mathbf{R}^p$. The states are $x(t) \in \mathbf{R}^n$, and the system has the form

$$\dot{x}(t) = f(x(t)) + g_w(x(t))w(t) + g_u(x(t))u(t), \quad (2)$$

where $f(0) = 0$. The objective is to solve the following problem.

Problem 1 Consider the nonlinear system (2). Construct a control law $\mu : \mathbf{R}^n \rightarrow \mathcal{U}$, with $\mu(0) = 0$, which is *robustly practically stabilizing* over a given set $\mathcal{X} \subseteq \mathbf{R}^n$, i.e., which guarantees that the state trajectory of the closed loop system

$$\dot{x}(t) = f(x(t)) + g_w(x(t))w(t) + g_u(x(t))\mu(x(t)) \quad (3)$$

will reach a positively invariant compact subset $\Omega \subset \mathcal{X}$ for all initial conditions $x(0) \in \mathcal{X}$, and all disturbance signals $w(t) \in \mathcal{W}$.

We remark here that the notion of robustness in this problem refers exclusively to the effect of the disturbance w ; in other words, the design is robust if stability is not compromised by the effect of the disturbance, as could happen in a nonlinear context. In contrast, the terminology of linear robust control refers to robustness with respect to unmodeled dynamics, where for example $w(t)$ would be a function of the state. We will discuss later a possible nonlinear extension of this more general notion of robustness. For the moment, we consider the special case introduced above, and adopt the following:

Definition 1 ([42]) We are given positive definite functions W and V . We say that V is a *robust control Lyapunov function (RCLF) with stability margin W* for the system (2) if there exist $c_2 > c_1 \geq 0$ and a control law $\mu : \mathbf{R}^n \rightarrow \mathcal{U}$

such that

$$\max_{w \in \mathcal{W}} L_f V(x) + L_{g_w} V(x)w + L_{g_u} V(x)\mu(x) + W(x) \leq 0$$

for all $x \in \mathbf{R}^n$ such that $c_1 \leq V(x) \leq c_2$. (Here and elsewhere, $L_f V$ stands for the directional derivative of V , in the direction of the vector field f .)

Note that a RCLF is decreasing along trajectories of the resulting closed loop system (3) whenever $w(t) \in \mathcal{W}$, and that this is a sufficient condition for robust practical stability [42, 112, 138].

If $\mathcal{U} = \mathbf{R}^m$, then Definition 1 is equivalent to the following.

Definition 2 V is a RCLF with stability margin W if and only if

$$\max_{w \in \mathcal{W}} L_f V(x) + L_{g_w} V(x)w + W(x) \leq 0$$

for all $x \in \mathbf{R}^n$ such that $c_1 \leq V(x) \leq c_2$ and $L_{g_u} V(x) = 0$.

Given a RCLF and its associated stability margin, a robustly practically stabilizing control law can be computed in a straightforward way using one of a number of *universal formulas* [3, 42, 74, 76, 114].

Recently, control Lyapunov functions have become an important topic of research. However, much of this research has concentrated on the properties of such functions [8, 24, 42, 43, 68, 75, 102, 120, 122, 123] with less attention given to their actual construction. Other references either assume that a Hamilton-Jacobi-Isaacs equation can be solved [5, 57, 58, 82, 83, 130, 131], or else consider special classes of nonlinear systems [20, 21, 41, 44, 66, 88, 105, 121].

In our research we have sought a solution to Problem 1 for a modified form of the class of systems considered in [8]; in particular, we treat systems of the form

$$\begin{bmatrix} \dot{x}_N \\ \dot{x}_L \end{bmatrix} = \begin{bmatrix} f_N(x_N) \\ f_L(x_N) \end{bmatrix} + \begin{bmatrix} A_N(x_N) \\ A_L(x_N) \end{bmatrix} x_L + \begin{bmatrix} g_w^N(x_N) \\ g_w^L(x_N) \end{bmatrix} w + \begin{bmatrix} 0 \\ g_u^L(x_N) \end{bmatrix} u \quad (4)$$

where the state vector x is partitioned into $x_N \in \mathbf{R}^k$ and $x_L \in \mathbf{R}^{n-k}$. This class of systems is most common in gain scheduling applications: for fixed values of the scheduling variables x_N (usually one or two-dimensional) the remaining dynamics are linear, and usually faster than the x_N dynamics. Therefore the x_N variables parametrize a “trim” surface around which one can control the system. Some recent literature on the gain scheduling problem includes [1, 2, 72, 100, 103, 106].

In the gain scheduling method, linear controllers are designed for the system linearized about trim points corresponding to various values of the scheduling variables, and some additional control logic is used to switch or interpolate between these controllers based on the values of the scheduling variables. A difficulty with this approach is the lack of any guarantee on the stability of the closed loop system. While some results on the stability *analysis* of such systems appear in the literature [107, 108], it is not clear how to *design* the gain scheduled controller for guaranteed stability.

We have developed an alternative solution to Problem 1, in which we select *robust control Lyapunov functions* for the system linearized about various trim points. At each point, we derive a RCLF (V_i, W_i) for the linearized dynamics using one of the available robust control techniques, depending on the nature of the performance objectives. Using the nonlinear dynamics, we compute the regions of attraction by computing the constants c_1, c_2 in the definition above. Notice that our objective is not to stabilize the system around the trim point, but rather move the trajectories around the trim surface to the equilibrium point. For this purpose, it is necessary that these regions intersect in a special way (we can describe this intersection in a precise way). The controller is then computed depending on the level set where the current state lies. Basically, the control action will switch forcing the trajectory to move from one region to the next until it reaches the equilibrium point, or an invariant set. A precise description of these results is given in [92, 91].

Our work has considered in detail the following issues associated with the above procedure:

1. There is the question of how to select the trim points, $V_i(x)$, and $W_i(x)$, in order to optimize performance. A preliminary answer to this question is to select $V_i(x)$ to be the quadratic Lyapunov function arising from the LQR or \mathcal{H}_∞ control design method applied to the linearized dynamics about a trim point. In addition, polytopic Lyapunov functions can be used in this procedure [90]. These capture a local peak-to-peak type performance.
2. While the literature contains many results for computing the stability region of an autonomous nonlinear system [18, 22, 31, 45, 50, 69, 94, 99, 111, 150], evaluating the stabilizability region requires the correct characterization of level sets of $V_i(x)$ intersected with the set $\ker(L_{g_u} V)$. Once this is accomplished, stabilizability can be analyzed for systems of the form (4) using results from nonconvex optimization theory [15, 16, 145, 146]. It is shown in [92] that computing the largest stability region associated with quadratic V_i and W_i is equivalent to a parametrized LMI, where the parameter is of the same dimension as x_N (typically low).
3. This procedure requires several iterations until the level sets of the RCLF cover the whole set \mathcal{X} . Methods for choosing the next trim point have been proposed. A complete analysis of the computational complexity of this procedure has been conducted [91].

2. Robust Stability for Nonlinear Operators

In the work of [59, 62], small gain results were derived for systems that are monotone stable. A nonlinear operator is stable in this sense if there exists a monotone function f such that

$$\|G(u)\| \leq f(\|u\|).$$

These results basically provided sufficient conditions for the stability in an input-output sense of the feedback interconnection of two such systems. If each system is associated with a function f_i , then stability is equivalent

(roughly) to the condition that the composition of these two functions is less than the identity map. In [46], we have analyzed the conservatism of this approach by showing classes of systems that will make this condition necessary. In particular, if one of the systems in the loop is a perturbation which is allowed to be any bounded causal operator, then it is shown that this condition is necessary (under some mild assumptions). This research parallels the development for LTI systems (see [29]).

3 Neuro-Dynamic Programming

3.1 Introduction

Dynamic programming (DP for short) provides a mathematical formalization and a general methodology for addressing problems in optimal stochastic control or sequential decision making under uncertainty [9]. The key construct in DP is the “cost-to-go” or “value” function, which is the total expected future cost, as a function of the initial state, under the assumption that an optimal policy (feedback control law) will be followed. Once this optimal value function is available, it can play the role of a control Lyapunov function and provides the basis for obtaining an optimal control law.

There are several numerical methods for computing the optimal value function, but they suffer generically from the “curse of dimensionality.” For example, the number of states in many important finite-state problems (Markov Decision Processes) is often overwhelming. Also, for continuous-time continuous-state problems, the DP approach leads to the Hamilton-Jacobi equation, which is very hard to solve numerically unless the dimension of the state is quite small. NDP attempts to overcome the curse of dimensionality by using a parametric representation of the value function (e.g., the value function can be represented in the form of a polynomial function of the state variables, or by a neural network with tunable weights). NDP methods involve on-line or simulation-based learning to tune the parameters of this parametric representation, in order to provide a sufficiently close approximation of the true value function, which then hopefully results in a close-to-optimal control law.

NDP has its origins in the AI community, under the name of “reinforcement learning” [67, 19], as well as in the early work by Werbös [143]. With exception of the pioneering work by Sutton [115] and Watkins [142], very little theory was available to support this methodology until the mid 1990s, even though the connections with dynamic programming were known. On the other hand, NDP methods have led to some remarkable successes (e.g., Tesauro’s world-class backgammon player [116]), which aroused a fair amount of interest on the subject. In the last 5 years or so, NDP theory has matured to a great extent, but the available results refer mostly to discrete time and/or discrete state problems [7, 10]. There are several methods that have been proposed in order to address continuous-state control problems described by nonlinear dynamics of the form that is common in control theory (see, e.g., the edited volume [144]). However, the available theory is mostly limited to linear quadratic regulator (LQR) problems, with a value function which is quadratic in the state variables [17, 70]. However, a quadratically parametrized value

function can exactly represent the true value function and, for this reason, the available results provide no insights on the behavior of such methods in the presence of approximation errors. Furthermore, the number of successful applications that have been reported and that involve continuous dynamics is rather limited [104, 51, 4]. For this reason, there is a clear need to develop the basic theory behind NDP for continuous control problems, and provide a streamlined and systematic design methodology. Furthermore, since the value function serves the role of a control Lyapunov function, there is also a need for a rapprochement between NDP and robust nonlinear control theory.

3.2 Summary of Past Research

1. Basic theory

As mentioned in the introduction, little theory existed to support NDP until the mid 1990s. Substantial progress has been made since then, linked to a significant extent to our work.

The best known and most popular NDP methods are based on Sutton's temporal difference (TD) algorithm [115] and Watkins' Q -learning algorithm [142]. These are simulation-based methods for learning the optimal value function or, sometimes, the value function associated with a fixed policy. Our first step was to explain these methods as stochastic approximation algorithms (of the Robbins-Monro type) for solving Bellman's equation, develop pertinent refinements of stochastic approximation theory, and provide convergence results. This was accomplished in our early work on the subject [124], which also refined and streamlined the available results on Q -learning, as well as in [60]. These results were restricted to the idealized case in which the algorithm maintains a numerical "estimate" of the true value function for each and every element of the state space. This case is of little practical interest, since it does not use approximations to overcome the curse of dimensionality, but is an important initial step.

Our subsequent work has dealt with the much more important case where the value function is parametrically represented. We solved a longstanding open problem by establishing the convergence of TD methods, for the case of a fixed control policy, as long as we are using a linearly parametrized approximation architecture, e.g. a linear combination of basis functions or "features" [126]. For more general (nonlinear) parametrizations, we have demonstrated that divergence is possible, even though this is rarely observed in practice. More important, we have developed error bounds that establish some basic "consistency" properties of TD methods: if the approximation architecture is rich enough to closely approximate the true value function, the the limit of TD will also provide a close approximation. These results are limited to discrete-time, infinite horizon, discounted cost problems, but encompass the case of a continuous or countably infinite state space. Subsequent work [128] proposed a TD method for problems involving an infinite horizon average cost criterion, and established similar results.

Recall that the above described results refer to the case where one estimates the value function associated with a single policy. They are of interest because such a "policy evaluation" forms the basis for policy improvement, which is how such methods are used in practice. Even though such a policy iteration approach is generically a

“discontinuous” algorithm, we have established its fundamental soundness by providing results with the following flavor: if each policy improvement is based on a sufficiently accurate approximate policy evaluation, then the resulting “approximate policy iteration” algorithm will eventually construct a policy which is close to optimal [10], even though the algorithm need not (and generically will not) converge.

In other work, we have identified two special cases for which TD or Q -learning methods are guaranteed to converge to a close to optimal policy (these are essentially the only available results of this type): (a) when the value function is approximated by a piecewise constant function [125]; (b) when we use a variant of Q -learning, together with a linearly parametrized approximation, to address optimal stopping problems [127].

2. Applications and case studies

The theory behind NDP methods can provide guidance for choosing promising approaches but not a guarantee for success. For this reason, we have found it important to apply several of these methods to a broad variety of interesting problems. The applications we have considered include job shop scheduling problems (where some NDP methods outperformed mainstream methods based on integer programming) [129, 23], scheduling in a reentrant line manufacturing system [139], a problem of machine maintenance and repair [11], an exotic options pricing problem [127], a problem of admission control and routing in an ATM communication network [86, 87], and a tracking problem for a missile with nonlinear dynamics [13].

Our experience with this experimentation has provided us with much insight into the “strong” and “weak” points of different methods, and into the proper choice of parametrized approximation architectures. It should be noted that with the exception of the missile control problem, all of our experiments have involved problems that can be formulated in discrete-time. Our experience with the missile control problem has indicated that a successful application of NDP to a continuous-time problem can be much more challenging.

3. A book on NDP

Together with Dimitri Bertsekas, we have written a book [10] on *Neuro-Dynamic Programming*. This book builds the foundations of the field, starting with relevant aspects of dynamic programming, and iterative learning theory. It is the only book on the subject that is available so far. It presents the state of the art at the theoretical front (including several results that have not been published elsewhere), as well as number of case studies. It has received the 1997 INFORMS Computer Science Technical Section prize for “research excellence at the interface between operations research and computer science.”

4 Identification of Complex Systems

4.1 Introduction

With the increasing awareness of the role of modeling errors on system performance has come a fresh look at the area of system identification based on data, motivated in part by the following:

1. In order for system identification tools to complement naturally those of robust control, a parallel development of this theory is required. In particular, the identified model should approximate the plant in a metric that is usable by robust control techniques.
2. The classical approach to system ID [80] assumes that data are generated by parametric models in the presence of stochastic noise. However, when using low-dimensional linear models for complex phenomena, a significant component of the error is due to deliberate under-modeling, rather than ambient noise. This distinction is of central importance for control systems, since unmodeled dynamics can be amplified by feedback and even lead to instabilities. For this reason, a system identification theory for control should allow for richer descriptions of uncertainty, in particular structured balls of unmodeled dynamics or nonlinearities. Equivalently, undermodeling should be explicitly incorporated in the problem formulation.
3. Even in regard to noise modeling, traditional system ID is restrictive since most of the theory applies only for stationary noise, modeled in a stochastic setting, and gives mostly *asymptotic* results. The results are not satisfactory for nonstationary noise, and little attention has been given to problems with finite data.
4. More generally, the understanding of the integrated picture of system identification and control design is still quite limited. Questions such as the fundamental limitations of system identification when the objective is to reduce the plant uncertainty, or the achievable performance of a control system when only finite corrupted data are available, are not well understood.

As a result, a research direction in identification for control has emerged and has attracted increasing interest in both the control and identification communities. In particular, the problem of identification for bounded (set-valued) noise has been extensively studied. The case where the objective is to optimize prediction for a fixed input was analyzed by many researchers [40, 81, 95, 96, 97, 98]. The problem is more interesting when the objective is to approximate the original system as an operator, a problem extensively discussed in [149]. For linear time invariant plants, such approximation can be achieved by uniformly approximating the frequency response (in the \mathcal{H}_∞ -norm) or the impulse response (in the ℓ_1 norm). In \mathcal{H}_∞ identification, it was shown that robustly convergent algorithms can be furnished, when the available data is in the form of a corrupted frequency response, at a set of points dense on the unit circle [52, 53, 54, 48, 49]. When the topology is induced by the ℓ_1 norm, a complete study of asymptotic identification was given in our past work [117, 118, 119, 27] for arbitrary inputs, and the question of optimal input design was addressed as well. Related work on this problem was also reported in [47, 61, 65, 71, 78, 79, 84, 85, 113].

Another issue of importance in the context of worst-case identification is *complexity*. It turns out that it is generally much harder to devise experiments that can guarantee small worst-case errors in the presence of bounded noise. This problem has been extensively analyzed in our work [28] and elsewhere [101, 77].

A related viewpoint, closely related with algorithm design, is model invalidation. In this case, one chooses a parametrization of a model including a description of the uncertainty as part of the parametrization, and the objective is to find a parameter which cannot be invalidated by a given set of finite data (see several articles in [113]). This is a weaker requirement since one only wants to find one out of the (possibly many) unfalsified models. If these models are to be used for control design, however, questions arise as to the size of the unfalsified set and conditions on the experiment to minimize the diameter of this set, which have not been addressed in the literature.

4.2 Summary of Past Research

We have proposed a new formulation for the system identification problem for complex systems. A space (\mathcal{T}) is complex if it cannot be uniformly approximated by a finite dimensional space. Nevertheless, we represent our *prejudice* by selecting a finitely parameterized set of models (\mathcal{G}) from which an estimate of the original system will ultimately be drawn. We will assume that an estimate of the distance (in some norm) between the actual process and this set is available as part of the prior information.

If the actual process is known, and if \mathcal{G} is convex, then selecting a model in \mathcal{G} that best approximates $T \in \mathcal{T}$ in some norm is a straightforward convex optimization problem. Hence, given any $T_0 \in \mathcal{T}$, we can write $T_0 = G_0 + \Delta_0$ where

$$G_0 = \arg \min_{G \in \mathcal{G}} \|T_0 - G\|$$

(for simplicity, assume the above minimization has a unique solution).

In system identification, however, the process is not known and only a finite set of input-output data is available. We will assume that this set of data is generated as:

$$y(k) = T_0 u(k) + w(k) = G_0 u(k) + \Delta_0 u(k) + w(k), \quad k \leq N$$

where u is the input (experiment) and w is a noise signal that belongs to some noise set. The objective of this development is to show, for rich classes of noise sets (either stochastic or deterministic that include white noise with high probability), how to select an input experiment u and an algorithm that picks an estimate $\hat{G}_N \in \mathcal{G}$ such that $\|G_0 - \hat{G}_N\|$ approaches zero in a reasonable length of time (hopefully with polynomial sample complexity). In other words, the derived algorithm used with the derived input provides a method for solving the actual approximation problem only from input-output data.

This work continues along the lines of [135] and distinguishes between the two sources of error — unmodeled dynamics and noise. We define a natural notion of separation between the parametric part and unmodeled dynamics. This notion arises naturally if the parametric part is a subspace of a linear space of systems, as discussed earlier. The noise is assumed to belong to a set of signals that are uncorrelated with the input (in either a deterministic

or stochastic sense) and is also rich enough to contain white noise sequences. Equipped with the triple — space of systems, parametric representation, and noise — we study consistency, sample complexity and algorithms. We emphasize that consistency is studied with respect to aforementioned decomposition. In particular, we study these issues for classes of systems consisting of ℓ_2 -stable systems and ℓ_∞ -stable systems.

For linear time-invariant systems, the story is complete in the following sense. Let \mathcal{T} be the space of all stable systems (this can be either ℓ_1 or H_∞), and \mathcal{G} be any finite dimensional subspace of finite-dimensional systems. Then it is possible to identify accurately (asymptotically) the element G_0 with sample complexity that is polynomial in the dimension of the subspace. In here G_0 is the best approximation of the system T_0 in \mathcal{G} . This requires the use of a special input (we term “robust” input) that has appropriate correlation properties that prove to be essential for obtaining consistent estimates in polynomial time. These results are reported in [136].

Instead of working in the ℓ_1 or H_∞ topologies, we can instead work with the Hardy-Sobolev topology, which gives rise to an inner product structure. The resulting norm provides an upper bound on the H_∞ norm. In this setting, it is shown that weighted least squares algorithms provide the appropriate estimates. The weights are constructed from the subspace \mathcal{G} in the form of non-causal filters that annihilate the contribution of \mathcal{G}^\perp . In addition, the algorithm is not tuned to the prior bound on the distance between the actual process and the subspace \mathcal{G} .

Identification in this setting provides estimates of the error in terms of the above norms. These estimates can consequently be used in a robust control setting for designing controllers. This is also true in the case where the norm is given by the Hardy-Sobolev norm even though it is not an induced measure. Results on robust control with such bounds can be found in [137].

The paradigm presented above bridges the gap between deterministic system identification formulations and stochastic ones. By appropriately selecting the noise set, one can give a deterministic description of white noise (or filtered white noise). This set is then used in the formulation of the system identification problem. The results are quite compatible with standard results (when exact modeling of the process is possible) in that one gets asymptotic convergence with polynomial sample complexity. The notion of persistence of excitation is also preserved. These results are reported in the Ph.D thesis [134].

5 Technology Transfer

During the course of past support, we had the following interactions with industry.

Robust Controller Design for the EOS-AM Model

Earth-Observing-Systems (EOS) are a cluster of small satellites that are intended to point at specific locations on earth, and obtain a variety of measurements (e.g. pictures of landscapes). The attitude control problem is quite important and design specifications were given (from *Draper Laboratory*) in terms of the peak to peak errors of attitude angles, in the presence of persistent disturbances and noise. The design should accommodate plant uncertainty, and take into account saturation and bandwidth constraints.

A controller was designed using the ℓ_1 toolbox. The toolbox handles both time and frequency domain specifications since it is based on computational methods (not closed form solutions). The results were compared to \mathcal{H}_∞ and μ designs as well as other classical methods, all of which were designed at Draper. We were able to show the limits of performance and draw tradeoff curves between optimal performance (measured in the time domain) and constraints (due to saturations and robustness). The solutions obtained were drastically different than the \mathcal{H}_∞ and μ designs, with a relative reduction of a factor of 50 in the Peak-to-Peak errors in the attitude angles. This highlights the advantages of a toolbox where these kinds of specifications can be put directly in the problem as opposed to obtaining them indirectly via frequency weight selection.

The results of these designs were reported in [39].

Active Vibration Isolation

Engineers at *Draper Laboratory* have been applying the ℓ_1 code to an active vibration isolation system, a Draper project demonstrating active structural control technology. Most of their emphasis has been on constraining the closed loop rather than optimization. This includes performance constraints, robustness constraints, and constraints on gain and phase margins. The resulting controller was tested on the actual hardware, and worked quite nicely. Reports on this are in progress [93].

Artillery Shells

The work on RCLF conducted in Mcconley's thesis was applied at *Draper Laboratory* to the control of artillery shells. The control strategy was tested on a complex nonlinear model for the shell and was shown to have very nice properties in terms of keeping the shell very close to the trim surface. This controller was much superior to a gain-scheduled controller which was shown to become unstable.

Feedback Control of OMVPE Growth

We have worked with *Spire Corporation* on the feedback control of OMVPE (Organo-metallic Vapor phase epitaxy) growth of compound semiconductor devices. The result of the work was a working controller (programmed in C) that controls both the thickness and the concentration to the specifications.

The controller was based on dynamic inversion, followed by a linear design. We used the ℓ_1 toolbox to quantify all the tradeoffs. Ultimately a much simpler design was implemented.

The results of this interaction can be found in [141, 140].

MIMO Acoustic Noise Cancellation

The research was conducted at *BBN (Bolt, Bernal and Newman)*, on two applications. One was on developing an adaptive chip for noise cancellation and the other was on developing a fixed control strategy for noise cancellation in a specific application for the Navy.

For the adaptive chip, the main constraint was memory availability and computational efficiency. For the system identification part, we demonstrated that the “instrumental variable method”, implemented recursively provides a computationally efficient method that does not require large memory availability. The controller design was based on classical design ideas.

For the second application, our contribution came in three folds. The first was showing that identification should be done in a behavioral setting using time series analysis. The second was showing how the Youla parametrization is essentially the rigorous treatment for acoustic feedback elimination. And the third was showing that the controller that minimizes the weighted \mathcal{H}_2 norm can be derived directly from data.

Hyperthermia Treatment of Prostate Cancer In the past two years, we started a collaborative effort with Brigham and Women’s on the problem of hyperthermia treatment of prostate cancer. Our preliminary work resulted in the design of an LQ-based controller that exhibited excellent distributed performance with minimal adaptation. The work is reported in [56]. This controller was implemented and tested. Currently, we are investigating the impact of adaptation on the design.

Neurophysiological and Neuroanatomical model of the Cerebellum

This work is currently being conducted in collaboration with NIH, through Prof. Dahleh’s Ph.D student Steve Massaquoi.

The brain of every animal that needs to generate quick accurate movements has a cerebellum, and a significant damage to the organ invariably results in clumsiness and in coordination. Drunken stagger, manual fumbling and slurring of speech are classic manifestations of cerebellar dysfunction. Because of the commonness of cerebellar disorders in humans, and the cerebellum’s relative accessibility to experimental investigations in animals, its physiology has been of great interest to neurologists and physiologists for some time. Much has been learned about its detailed neuroanatomy and cellular neurophysiology.

The objective of this research is to build a neuroanatomically and neurophysiologically consistent (intelligent) control theoretic model for the cerebellum. Our interest is motivated by several factors: 1. the relative availability of physiological data, 2. the connections between such a model and the parallel neural networks architectures, 3. the hybrid nature of the control system, and 4. the lack of a control oriented model that captures the different features

of the cerebellum.

Through preliminary investigations, we have shown that velocity feedback results in a system that accurately described the data collected from people with both healthy and defective cerebellums. The difference between the two is primarily the gain. In addition, we have given a wave variable interpretation for the delay compensation which borrows from other people's work in the robotics area. Many of these results have been reported in [89].

Development of the ℓ_1 toolbox

While the above applications were primarily performed by us, the software has been made available to various corporations, laboratories and universities.

Multi-Asset Management and NDP

In research performed at Alphatech, Inc., we collaborated in the development of novel approaches for multi-asset management in surveillance by UAVs, a very complex problem involving path generation, assignment of tasks to sensors, and scheduling. Our methodology involves a combination of mathematical programming approaches, together with NDP methods for track generation. As a consequence of this work, NDP is now listed in DARPA's Advanced Cooperative Collection Management Web pages as one of few possible enabling technologies. Our approach appears very promising for a broad class of vehicle routing problems with time windows (VRPTW)

NDP for supply-chain management

In research performed at Unica Technologies, Inc., we participated in a project meant to demonstrate the promise of NDP methods for supply-chain management (multi-echelon inventory control), with the goal of transferring this technology to practical use. The results, reported in [132] have been positive, as NDP methods outperformed an optimized heuristic on a difficult problem involving more than 40 state variables.

NDP in ATM networks

In collaboration with Siemens, we have been developing NDP methods for admission control and routing in ATM networks, in the presence of multiple customer classes with different values. A particular type of a decomposable architecture has been developed to suit this kind of problem, and encouraging results have been reported in [86, 87].

Interactions with Wright-Patterson Labs.

Prof. Tsitsiklis visited the group of Dr. Klopff at WPAFB in the fall of 1996, and had extensive discussions on the subject of NDP and reinforcement learning. There was agreement on the promise of such methods in the area of sensor scheduling, sensor management, and sensor fusion, which is of great importance for the Air Force. As a followup to this visit, Lt. Harmon from WPAFB visited MIT in November of 1996, made a presentation on interactive modular software, available on the Web, that can support NDP experiments, and interacted with members of our research group.

6 Major Publications

1. D. Bertsekas and J. Tsitsiklis. *Neuro-dynamic programming*, Athena Scientific, Belmont, MA, 1996.
2. D. P. Bertsekas, J. N. Tsitsiklis, and C. Wu, "Rollout Algorithms for Combinatorial Optimization", to appear in the *Journal of Heuristics*.
3. M.A. Dahleh and I. Diaz-Bobillo, *Control of uncertain systems: A linear programming approach*, Prentice-Hall, N.J. 1995.
4. N. Elia and M.A. Dahleh. *Multi-objective controller design*, Springer Verlag. To appear.
5. N. Elia and M.A. Dahleh, "Controller Design with Multiple Objectives", *IEEE Trans. Automat. Contr.*, Vol 42, No. 5, pp. 596-613, May 1997.
6. Magnitude Constraints in the Frequency Domain", to appear in the Journal of Optimization Theory and Applications.
7. N. Elia and M.A. Dahleh, "Minimization of the Worst-Case Peak to Peak Gain via Dynamic Programming". Accepted for publication in *IEEE Trans. A-C*.
8. J. Goncalves and M.A. Dahleh. "Necessary and sufficient conditions for robust stability of a class of nonlinear systems," *Automatica*, Vol 34, No. 6, pp. 705-714, 1998.
9. P. Marbach, O. Mihatsch, O., M. Schulte, and J. N. Tsitsiklis", "Reinforcement Learning for Call Admission Control and Routing in Integrated Service Networks", in *Proc. 1997 NIPS*, Denver, Colorado.
10. P. Marbach and J.N.Tsitsiklis, "A Neuro-Dynamic Programming Approach to Admission Control in ATM Networks: The Single Link Case", International Conference on Acoustic, Speech, and Signal Processing, Muenchen, Germany, April 1997.
11. M.W. McConley, M.A. Dahleh and E. Feron. "Polytopic Control Lyapunov Functions for Robust Stabilization of a Class of Nonlinear Systems," *Systems and Control Letters*, 34, pp. 77-83, 1998.
12. M.W. McConley, B.D. Appleby, M.A. "Dahleh and E. Feron. "Computational Complexity of Lyapunov Stability Analysis Problems for a Class of Nonlinear Systems," accepted for publication in the *SIAM* journal on Control and Optimization.
13. M.W. McConley, B.D. Appleby, M.A. "Dahleh and E. Feron. A Control Lyapunov Function Approach to Robust Stabilization of Nonlinear Systems," accepted for publication in *IEEE Trans. A-C*.

14. J. N. Tsitsiklis and B. Van Roy, "Feature-Based Methods for Large Scale Dynamic Programming", *Machine Learning*, Vol. 22, 1996, pp. 59-94.
15. J. N. Tsitsiklis and B. Van Roy, "An Analysis of Temporal-Difference Learning with Function Approximation", *IEEE Transactions on Automatic Control*, Vol. 42, No. 5, May 1997, pp. 674-690.
16. J. N. Tsitsiklis, and B. Van Roy, "Optimal Stopping of Markov Processes: Hilbert Space Theory, Approximation Algorithms, and an Application to Pricing Financial Derivatives", submitted to the *IEEE Transactions on Automatic Control*, 1997.
17. J. N. Tsitsiklis, and B. Van Roy, "Average Cost Temporal-Difference Learning", submitted to *Automatica*, 1997.
18. J. N. Tsitsiklis and J. D. Christodouleas, "Approximate DP Methods for Multiprocessor Network Scheduling", in preparation.
19. S. Venkatesh and M.A. Dahleh. "Identification in the presence of classes of unmodeled dynamics and noise," *IEEE Trans. A-C*, December 1997.
20. S. Venkatesh and M.A. Dahleh, "Identification of Complex systems with limited-complexity models," Submitted to *IEEE Trans. A-C*.
21. S. Venkatesh, A. Megretski and M.A. Dahleh, "A convex parametrization of stabilizing controllers for perturbations belonging to a Hardy-Sobolev Space," submitted.
22. S. Warnick, M.A. Dahleh, "Feedback control of OMVPE growth of submicron compound semiconductor films", *IEEE Trans. on Control Systems Technology*, Vol 6, pp 62-71, Jan 1998.

7 Theses Supervised

Several theses have been supervised with partial support from this grant.

1. M. Escobar. "Systematic procedure to meet specific input-output constraints in ℓ_1 optimal control problem design," LIDS thesis, Jan, 1995. Supervised by Prof. M.A. Dahleh.
2. S. Warnick. "Feedback control of OMVPE growth of compound semiconductor devices", LIDS thesis, August 1995, supervised by Prof. M.A. Dahleh.
3. S. Venkatesh. Identification for Complex Systems. MIT Ph.D thesis No. LIDS-Th-2394, July 1997, supervised by Prof. Dahleh.

4. C. Boussios. An approach for nonlinear control design via approximate dynamic programming, doctoral thesis, MIT, May 1998, supervised by Profs. M.A. Dahleh and J.N. Tsitsiklis.
5. M. McConley. A computationally efficient Lyapunov-based procedure for control of nonlinear systems with stability and performance guarantees, LIDS PhD thesis, May 1997, supervised by Prof. Dahleh.

References

- [1] P. Apkarian and P. Gahinet. A convex characterization of gain-scheduled H_∞ controllers. *IEEE Transactions on Automatic Control*, 40(5):853–864, May 1995.
- [2] P. Apkarian, P. Gahinet, and G. Becker. Self-scheduled H_∞ control of linear parameter-varying systems. In *Proceedings of the American Control Conference*, pages 856–860, Baltimore, MD, June 1994.
- [3] Z. Artstein. Stabilization with relaxed controls. *Nonlinear Analysis, Theory, Methods & Applications*, 7(11):1163–1173, 1983.
- [4] S. N. Balakrishnan, and V. Biega, “A New Neural Architecture for Homing Missile Guidance”, *Proc. American Control Conference*, June 1995, pp. 2148-2152.
- [5] J. A. Ball, W. Helton, and M. L. Walker. H^∞ control for nonlinear systems with output feedback. *IEEE Transactions on Automatic Control*, 38(4):546–559, Apr. 1993.
- [6] A.E. Barabanov, and A.A. Sokolov, “A Geometrical Solution to ℓ_1 optimization with combined conditions.” In *Proceedings of Asian Control Conference*, Tokyo Japan, vol. 3, pp 331-334, 1994.
- [7] Barto, A. G., Bradtke, S. J., and Singh, S. P., “Learning to Act Using Real-Time Dynamic Programming,” *Artificial Intelligence*, Vol. 72, pp. 81-138, 1995.
- [8] S. Battilotti. Stabilization via dynamic output feedback for systems with output nonlinearities. *Systems & Control Letters*, 23(6):411–419, Dec. 1994.
- [9] Bertsekas, D. P., *Dynamic Programming and Optimal Control*, Vols. I and II, Athena Scientific, Belmont, MA, 1995.
- [10] D. Bertsekas and J. Tsitsiklis. *Neuro-dynamic programming*, Athena Scientific, Belmont, MA, 1996.
- [11] D. P. Bertsekas, J. N. Tsitsiklis, and C. Wu, “Rollout Algorithms for Combinatorial Optimization”, to appear in the *Journal of Heuristics*.
- [12] F. Blanchini, and M. Sznaier, “Persistent Disturbance Rejection via Static Feedback” *IEEE Transactions on Automatic Control*. 40, 1127-1131, 1995.
- [13] C. Boussios, An approach for nonlinear control design via approximate dynamic programming, doctoral thesis, MIT, May 1998, supervised by Profs. M.A. Dahleh and J.N. Tsitsiklis.
- [14] S. Boyd and W. Barrat. *Linear controller design: Limits of performance*, Prentice-Hall, NJ, 1991.

- [15] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 1994.
- [16] R. P. Braatz, P. M. Young, J. C. Doyle, and M. Morari. Computational complexity of μ calculation. *IEEE Transactions on Automatic Control*, 39(5):1000–1002, May 1994.
- [17] Bradtke, S. J., Incremental Dynamic Programming for On-Line Adaptive Optimal Control, Ph. D. thesis, University of Massachusetts, Amherst, MA, 1994.
- [18] R. K. Brayton and C. H. Tong. Constructive stability and asymptotic stability of dynamical systems. *IEEE Transactions on Circuits and Systems*, 27(11):1121–1130, Nov. 1980.
- [19] Barto, A. G., Sutton, R. S., and Anderson, C. W., “Neuronlike Elements that Can Solve Difficult Learning Control Problems,” *IEEE Trans. on Systems, Man, and Cybernetics*, Vol. 13, pp. 835–846, 1983.
- [20] C. I. Byrnes and A. Isidori. New results and examples in nonlinear feedback stabilization. *Systems & Control Letters*, 12(5):437–442, June 1989.
- [21] C. I. Byrnes and A. Isidori. Asymptotic stabilization of minimum phase nonlinear systems. *IEEE Transactions on Automatic Control*, 36(10):1122–1137, Oct. 1991.
- [22] H.-D. Chiang and J. S. Thorp. Stability regions of nonlinear dynamical systems: A constructive methodology. *IEEE Transactions on Automatic Control*, 34(12):1229–1241, Dec. 1989.
- [23] J. Christodouleas, “Solution Methods for Multiprocessor Network Scheduling Problems, with Applications to Railroad Operations”, doctoral thesis, MIT, 1997, supervised by Prof. J.N. Tsitsiklis.
- [24] J.-M. Coron and L. Praly. Adding an integrator for the stabilization problem. *Systems & Control Letters*, 17:89–104, 1991.
- [25] M.A. Dahleh and M.H. Khammash, “Controller Design in the Presence of Structured Uncertainty,” *Automatica*, Vol 29, No. 1, pp. 37–56, January 1993.
- [26] M.A. Dahleh and Y. Ohta. “A necessary and sufficient condition for robust BIBO stability,” *Systems and Control Letters*, Vol. 11, pp. 271–275, 1988.
- [27] M.A. Dahleh, E.D. Sontag, D.N. Tse and J.N. Tsitsiklis, “Worst-Case identification of nonlinear fading memory systems,” *Automatica*, Vol. 31, No. 3, March 1995, pp. 503–508.
- [28] M.A. Dahleh, T. Theodosopoulos, and J.N. Tsitsiklis, “The sample complexity of worst-case identification of F.I.R. Linear systems,” *System and Control. letters*, Vol 20, No. 3, pp. 157–166, March 1993.

- [29] M.A. Dahleh and I. Diaz-Bobillo, *Control of uncertain systems: A linear programming approach*, Prentice-Hall, N.J. 1995.
- [30] J. C. Doyle, "Analysis of feedback systems with structured uncertainty," *IEEE Proceedings* 129, 242-250, 1982.
- [31] E. J. Davison and E. M. Kurak. A computational method for determining quadratic Lyapunov functions for non-linear systems. *Automatica*, 7:627-636, 1971.
- [32] I. Diaz-Bobillo and M.A. Dahleh, "State Feedback ℓ_1 Optimal Controllers Can be Dynamic," *System and Control Letters*, Vol 19, No. 2, pp. 87-93, August 1992.
- [33] I. Diaz-Bobillo and M.A. Dahleh. "Minimization of the maximum peak-to-peak gain: The General multiblock problem," *IEEE Trans. Automat. Contr.*, Vol 38, No. 10, pp. 1459-1482, October 1993.
- [34] N. Elia, M.A. Dahleh and I.J. Diaz-Bobillo. "Controller design via linear programming," *Proc ACC*, San Francisco, CA, June 1993, pp. 2165-2169.
- [35] N. Elia and M.A. Dahleh, "Controller Design with Multiple Objectives", *IEEE Trans. Automat. Contr.*, Vol 42, No. 5, pp. 596-613, May 1997.
- [36] N. Elia and M. A. Dahleh. " ℓ_1 Minimization with Magnitude Constraints in the Frequency Domain", to appear in the *Journal of Optimization Theory and Applications*.
- [37] N. Elia, P. Young and M.A. Dahleh. "Robust Performance with fixed inputs", *Proc. CDC*, Lake Buena Vista, FL, 1994, pp. 2690-2695.
- [38] N. Elia and M.A. Dahleh, "Minimization of the Worst-Case Peak to Peak Gain via Dynamic Programming". Accepted for publication in *IEEE Trans. A-C*.
- [39] M. Escobar. "Systematic procedure to meet specific input-output constraints in ℓ_1 optimal control problem design," LIDS thesis, Jan, 1995. Supervised by Prof. M.A. Dahleh.
- [40] E. Fogel and Y. F. Huang, "On the value of information in system identification-bounded noise case", *Automatica*, vol.18, no.2, pp.229-238, 1982.
- [41] R. A. Freeman and P. V. Kokotović. Design of 'softer' robust nonlinear control laws. *Automatica*, 29(6):1425-1437, 1993.
- [42] R. A. Freeman and P. V. Kokotović. Inverse optimality in robust stabilization. To appear, Oct. 1994.

- [43] R. A. Freeman and P. V. Kokotović. Robust control Lyapunov functions: The measurement feedback case. In *Proceedings of the IEEE Conference on Decision and Control*, pages 3533–3538, Lake Buena Vista, FL, Dec. 1994.
- [44] R. A. Freeman and P. V. Kokotović. Tools and procedures for robust control of nonlinear systems. In *Proceedings of the IEEE Conference on Decision and Control*, pages 3458–3463, Lake Buena Vista, FL, Dec. 1994.
- [45] R. Genesio, M. Tartaglia, and A. Vicino. On the estimation of asymptotic stability regions: State of the art and new proposals. *IEEE Transactions on Automatic Control*, 30(8):747–755, Aug. 1985.
- [46] J. Goncalves and M.A. Dahleh. “Necessary and sufficient conditions for robust stability of a class of nonlinear systems,” *Proc. CDC*, New Orleans, Louisiana, 1995.
- [47] G.C. Goodwin. “Experiment design in system identification” in *Encyclopedia of Systems and Control* (M. Singh, Ed.), Pergamon Press, 1987.
- [48] G. Gu, P.P. Khargonekar and Y. Li, “Robust convergence of two stage nonlinear algorithms for identification in \mathcal{H}_∞ ,” *Systems and Control Letters*, Vol 18, No. 4, April 1992.
- [49] G. Gu and P.P. Khargonekar, “Linear and nonlinear algorithms for identification in \mathcal{H}_∞ with error bounds,” *IEEE Trans. A-C*, Vol 37, No. 7, July 1992.
- [50] W. Hahn. *Stability of Motion*. Springer-Verlag, New York, 1967.
- [51] M. E. Harmon, L. C. Baird III, A. H. Klopff, “Reinforcement Learning Applied to a Differential Game”, *Adaptive Behavior*, Vol. 4, No. 1, pp. 3-28, 1996.
- [52] A.J. Helmicki, C.A. Jacobson and C.N. Nett, “Identification in H_∞ : A robust convergent nonlinear algorithm”, *Proceedings of the 1989 International Symposium on the Mathematical Theory of Networks and System*, 1989.
- [53] A.J. Helmicki, C.A. Jacobson and C.N. Nett, “Identification in H_∞ : Linear Algorithms”, *Proceedings of the 1990 American Control Conference*, pp 2418-2423.
- [54] A.J. Helmicki, C.A. Jacobson and C.N. Nett, “Control-oriented System Identification: A Worst-case/deterministic Approach in H_∞ ,” *IEEE Trans. A-C*, Vol 36, No. 10, October 1991.
- [55] A. Isidori. *Nonlinear Control Systems: An Introduction*, Springer Verlag, 1989.
- [56] E. Hutchinson, M.A. Dahleh and K. Hynynen, “The feasibility of MRI Feedback Control for Intracavitary Phased Array Hyperthermia Treatments”, accepted for publications in *International Journal of Hyperthermia*.

- [57] A. Isidori and A. Astolfi. Disturbance attenuation and H_∞ -control via measurement feedback in nonlinear systems. *IEEE Transactions on Automatic Control*, 37(9):1283–1293, Sept. 1992.
- [58] A. Isidori and W. Kang. H_∞ control via measurement feedback for general nonlinear systems. *IEEE Transactions on Automatic Control*, 40(3):466–472, Mar. 1995.
- [59] M. Ives, Y. Mareels and D. Hill, "Monotone stability of nonlinear feedback systems", *Journal of Mathematical Systems, Estimation and Control*, 2(3): 275-291, 1992.
- [60] Jaakkola, T., Jordan, M. I., and Singh, S. P., "On the Convergence of Stochastic Iterative Dynamic Programming Algorithms," *Neural Computation*, Vol. 6, pp. 1185-1201, 1994.
- [61] C.A. Jacobson and C.N. Nett, "Worst-case system identification in ℓ_1 : Optimal algorithms and error bounds," in *Proc. of the 1991 American Control Conference*, June 1991.
- [62] Z. Jinag, A. Teel and Praly, "Small gain theorem for ISS systems and application", *MCSS*, 7: 95-120, 1994.
- [63] M. Khammash, "Robust steady-state tracking", to appear in *IEEE Trans, A-C*.
- [64] M.H. Khammash and J.B. Pearson. "Stability and performance robustness in l^1 optimal control systems," Rice University *Technical Report*, Houston, TX.
- [65] J.M. Krause. *Robust identification and adaptation*, Ph.D Thesis, University of Minnesota, 1987.
- [66] I. Kanellakopoulos, P. V. Kokotović, and A. S. Morse. A toolkit for nonlinear feedback design. *Systems & Control Letters*, 18(2):83–92, Feb. 1992.
- [67] A. H. Klopff, *The Hedonistic Neuron: A Theory of Memory, Learning, and Intelligence*, Hemisphere, Washington, DC, 1982.
- [68] M. Krstić and P. V. Kokotović. Control Lyapunov functions for adaptive nonlinear stabilization. *Systems & Control Letters*, 26(1):17–23, Sept. 1995.
- [69] M. C. Lai and J. Hauser. Computing maximal stability region using a given Lyapunov function. In *Proceedings of the American Control Conference*, pages 1500–1502, San Francisco, June 1993.
- [70] T. Landelius, Reinforcement Learning and Distributed Local Model Synthesis, Ph.D. thesis, Linköping University, Sweden, 1997.
- [71] M.K. Lau, R.L. Kosut, S. Boyd, "Parameter Set Estimation of Systems with Uncertain Nonparametric Dynamics and Disturbances", *Proceedings of the 29th Conference on Decision and Control*, pp. 3162-3167, 1990.

- [72] D. A. Lawrence and W. J. Rugh. Gain scheduling dynamic linear controllers for a nonlinear plant. *Automatica*, 31(3):381–390, Mar. 1995.
- [73] The Control Handbook. CRC Press, 1996. Edited by W. Levine.
- [74] W. Lin. Bounded smooth state feedback and a global separation principle for non-affine nonlinear systems. *Systems & Control Letters*, 26(1):41–53, Sept. 1995.
- [75] Y. Lin, E. Sontag, and Y. Wang. Recent results on Lyapunov-theoretic techniques for nonlinear stability. In *Proceedings of the American Control Conference*, pages 1771–1775, Baltimore, MD, June 1994.
- [76] Y. Lin and E. D. Sontag. A universal formula for stabilization with bounded controls. *Systems & Control Letters*, 16(6):393–397, June 1991.
- [77] L. Lin, L. Wang and G. Zames, “Uncertainty principle and identification n-Widths for LTI and slowly varying systems,” *ACC*, Chicago, IL, 1992.
- [78] M. Livstone and M.A. Dahleh. “A framework for robust parametric set membership identification,” *IEEE Trans. A-C*, vol 40, pp. 1934–1939, Nov 1995.
- [79] M. Livstone and M.A. Dahleh. “Set membership identification algorithms and asymptotic properties”, to appear in *Systems and Control Letters*.
- [80] L. Ljung. *System identification: theory for the user*, Prentice-Hall, Inc, NJ, 1987.
- [81] R. Lozano-Leal and R. Ortega, “Reformulation of the parameter identification problem for systems with bounded disturbances”, *Automatica*, vol.23, no.2, pp.247–251, 1987.
- [82] W.-M. Lu and J. C. Doyle. H_∞ control of nonlinear systems via output feedback: Controller parameterization. *IEEE Transactions on Automatic Control*, 39(12):2517–2521, Dec. 1994.
- [83] W.-M. Lu and J. C. Doyle. A state space approach to robustness analysis and synthesis for nonlinear uncertain systems. Technical Report 94-146, California Institute of Technology, Pasadena, CA, Mar. 1994.
- [84] P.M. Makila, “Robust Identification and Galois Sequences”, Technical Report 91-1, Process Control Laboratory, Swedish University of Abo, January, 1991.
- [85] P.M. Makila and J.R. Partington, “Robust Approximation and Identification in H_∞ ”, Proc. 1991 American Control Conference, June, 1991.
- [86] P. Marbach, O. Mihatsch, O., M. Schulte, and J. N. Tsitsiklis, “Reinforcement Learning for Call Admission Control and Routing in Integrated Service Networks”, in *Proc. 1997 NIPS*, Denver, Colorado.

- [87] P. Marbach and J.N.Tsitsiklis, "A Neuro-Dynamic Programming Approach to Admission Control in ATM Networks: The Single Link Case", International Conference on Acoustic, Speech, and Signal Processing, Muenchen, Germany, April 1997.
- [88] R. Marino and P. Tomei. Robust stabilization of feedback linearizable time-varying uncertain nonlinear systems. *Automatica*, 29(1):181-189, 1993.
- [89] S. Massaquoi. "Evaluation of Linear Multiinput-Multioutput Feedback Control for a Human Arm Control", LIDS-TH-2198. Supervised by Prof. M.A. Dahleh.
- [90] M.W. McConley, M.A. Dahleh and E. Feron. "Polytopic Control Lyapunov Functions for Robust Stabilization of a Class of Nonlinear Systems," *Systems and Control Letters*, 34, pp. 77-83, 1998.
- [91] M.W. McConley, B.D. Appleby, M.A. "Dahleh and E. Feron. "Computational Complexity of Lyapunov Stability Analysis Problems for a Class of Nonlinear Systems," accepted for publication in the *SIAM* journal on Control and Optimization.
- [92] M.W. McConley, B.D. Appleby, M.A. "Dahleh and E. Feron. A Control Lyapunov Function Approach to Robust Stabilization of Nonlinear Systems," accepted for publication in *IEEE Trans. A-C*.
- [93] L. McGovern, "Application of the ℓ_1 robust control methodology in active vibration isolation," S.M. thesis, June 1996.
- [94] A. N. Michel, N. R. Sarabudla, and R. K. Miller. Stability analysis of complex dynamical systems. *Circuits, Systems, and Signal Processing*, 1(2):171-202, 1982.
- [95] M. Milanese and G. Belforte, "Estimation theory and uncertainty intervals evaluation in the presence of unknown but bounded errors: Linear families of models and estimators", *IEEE Trans. Automatic Control*, AC-27, pp.408-414, 1982.
- [96] M. Milanese and R. Tempo, "Optimal algorithm theory for robust estimation and prediction", *IEEE Trans. Automatic Control*, AC-30, pp. 730-738, 1985.
- [97] M. Milanese, "Estimation theory and prediction in the presence of unknown and bounded uncertainty: a survey", in *Robustness in Identification and Control*, M. Milanese, R. Tempo, A. Vicino Eds, Plenum Press, 1989.
- [98] J.P. Norton, "Identification and application of bounded-parameter models", *Automatica*, vol.23, no.4, pp.497-507, 1987.
- [99] Y. Ohta et al. Computer generated Lyapunov functions for a class of nonlinear systems. *IEEE Transactions on Circuits and Systems — I: Fundamental Theory and Applications*, 40(5):343-354, May 1993.

- [100] A. Packard. Gain scheduling via linear fractional transformations. *Systems & Control Letters*, 22(2):79–92, Feb. 1994.
- [101] K. Poolla and A. Tikku, “On the time complexity of worst-case system identification,” *IEEE Trans. AC.*, May 1994.
- [102] L. Praly, B. d’Andréa Novel, and J.-M. Coron. Lyapunov design of stabilizing controllers for cascaded systems. *IEEE Transactions on Automatic Control*, 36(10):1177–1181, Oct. 1991.
- [103] W. J. Rugh. Analytical framework for gain scheduling. *IEEE Control Systems Magazine*, 11(1):79–84, Jan. 1991.
- [104] Prokhorov, D. V., Santiago, R. A., and Wunsch, D. C., II, “Adaptive Critic Designs: A Case Study for Neurocontrol,” *Automatica*, Vol. 8, pp. 1367–1372, 1995.
- [105] D. Seto, A. M. Annaswamy, and J. Baillieul. Adaptive control of nonlinear systems with a triangular structure. *IEEE Transactions on Automatic Control*, 39(7):1411–1428, July 1994.
- [106] S. M. Shahrzad and S. Behtash. Design of controllers for linear parameter-varying systems by the gain scheduling technique. *Journal of Mathematical Analysis and Applications*, 168(1):195–217, July 1992.
- [107] J. S. Shamma and M. Athans. Analysis of gain scheduled control for nonlinear plants. *IEEE Transactions on Automatic Control*, 35(8):898–907, Aug. 1990.
- [108] J. S. Shamma and M. Athans. Guaranteed properties of gain scheduled control for linear parameter-varying plants. *Automatica*, 27(3):559–564, May 1991.
- [109] J.S. Shamma, “Nonlinear state feedback for ℓ_1 optimal control”, *Systems and Control Letters*, October 1993, pp. 265–270.
- [110] J.S. Shamma, “Optimization of the ℓ_∞ induced norm under full state feedback”, *IEEE Transactions on Automatic Control*, April 1996.
- [111] D. N. Shields and C. Storey. The behaviour of optimal Lyapunov functions. *International Journal of Control*, 21(4):561–573, 1975.
- [112] J.-J. E. Slotine and W. Li. *Applied Nonlinear Control*. Prentice-Hall, Inc., Englewood Cliffs, NJ, 1991.
- [113] R. Smith and M. Dahleh. *The Modeling of Uncertainty in Control Systems*, Springer-Verlag, 1994.
- [114] E. D. Sontag. A ‘universal’ construction of Artstein’s theorem on nonlinear stabilization. *Systems & Control Letters*, 13(2):117–123, Aug. 1989.

- [115] Sutton, R. S., "Learning to Predict by the Methods of Temporal Differences," *Machine Learning*, Vol. 3, pp. 9-44, 1988.
- [116] Tesauro, G. J., "TD-Gammon, a Self-Teaching Backgammon Program, Achieves Master-Level Play," *Neural Computation*, Vol. 6, pp. 215-219, 1994.
- [117] D. Tse, M.A. Dahleh and J.N. Tsitsiklis. Optimal Asymptotic Identification under bounded disturbances. *IEEE Trans. Automat. Contr.*, Vol. 38, No. 8, pp. 1176-1190, August 1993.
- [118] D.N.C. Tse, M.A. Dahleh, J.N. Tsitsiklis, "Optimal and Robust Identification in the ℓ_1 norm", in *Proc. of the 1991 American Control Conference*, June 1991.
- [119] D.N.C. Tse, "Optimal and robust identification under bounded disturbances", Master's thesis, Dept. of Elec. Eng. and Comp. Sci., M.I.T., February, 1991.
- [120] J. Tsiniias. Existence of control Lyapunov functions and applications to state feedback stabilizability of nonlinear systems. *SIAM Journal on Control and Optimization*, 29(2):457-473, Mar. 1991.
- [121] J. Tsiniias. A theorem on global stabilization of nonlinear systems by linear feedback. *Systems & Control Letters*, 17(5):357-362, Nov. 1991.
- [122] J. Tsiniias. On the existence of control Lyapunov functions: Generalizations of Vidyasagar's theorem on nonlinear stabilization. *SIAM Journal on Control and Optimization*, 30(4):879-893, July 1992.
- [123] J. Tsiniias. Versions of Sontag's "input to state stability condition" and the global stabilizability problem. *SIAM Journal on Control and Optimization*, 31(4):928-941, July 1993.
- [124] J. N. Tsitsiklis, "Asynchronous Stochastic Approximation and Q-learning", *Machine Learning*, 16, 1994, pp. 185-202.
- [125] J. N. Tsitsiklis and B. Van Roy, "Feature-Based Methods for Large Scale Dynamic Programming", *Machine Learning*, Vol. 22, 1996, pp. 59-94.
- [126] J. N. Tsitsiklis and B. Van Roy, "An Analysis of Temporal-Difference Learning with Function Approximation", *IEEE Transactions on Automatic Control*, Vol. 42, No. 5, May 1997, pp. 674-690.
- [127] J. N. Tsitsiklis, and B. Van Roy, "Optimal Stopping of Markov Processes: Hilbert Space Theory, Approximation Algorithms, and an Application to Pricing Financial Derivatives", submitted to the *IEEE Transactions on Automatic Control*, 1997.

- [128] J. N. Tsitsiklis, and B. Van Roy, "Average Cost Temporal-Difference Learning", submitted to *Automatica*, 1997.
- [129] J. N. Tsitsiklis and J. D. Christodouleas, "Approximate DP Methods for Multiprocessor Network Scheduling", in preparation.
- [130] A. J. van der Schaft. On a state space approach to nonlinear H_∞ control. *Systems & Control Letters*, 16:1-8, 1991.
- [131] A. J. van der Schaft. L_2 -gain analysis of nonlinear systems and nonlinear state feedback H_∞ control. *IEEE Transactions on Automatic Control*, 37(6):770-784, June 1992.
- [132] B. Van Roy, D. P. Bertsekas, Y. Lee, and J. N. Tsitsiklis, "A Neuro-Dynamic Programming Approach to Retailer Inventory Management", submitted to the *IEEE Trans. on Neural Networks*, 1997.
- [133] S.R. Venkatesh, "Properties of ℓ_1 optimal solutions," MIT S.M. Thesis, February 1993.
- [134] S. Venkatesh. Identification for Complex Systems. MIT Ph.D thesis No. LIDS-Th-2394, July 1997.
- [135]
- [136] S. Venkatesh and M.A. Dahleh. "Identification in the presence of classes of unmodeled dynamics and noise," *IEEE Trans. A-C*, December 1997.
- [137] S. Venkatesh and M.A. Dahleh, "Identification of Complex systems with limited-complexity models," Submitted to *IEEE Trans. A-C*.
- [138] S. Venkatesh, A. Megretski and M.A. Dahleh, "A convex parametrization of stabilizing controllers for perturbations belonging to a Hardy-Sobolev Space," submitted.
- [139] M. Vidyasagar. *Nonlinear Systems Analysis*. Prentice-Hall, Inc., Englewood Cliffs, NJ, second edition, 1993.
- [140] Vyzas, E., "Approximate Dynamic Programming for some Queueing Problems", Masters thesis, February 1997, supervised by Prof. J.N. Tsitsiklis.
- [141] S. Warnick. "Feedback control of OMVPE growth of compound semiconductor devices", LIDS thesis, August 1995. Supervised by Prof. M.A. Dahleh.
- [142] S. Warnick, M.A. Dahleh, "Feedback control of OMVPE growth of submicron compound semiconductor films", *IEEE Trans. on Control Systems Technology*, Vol 6, pp 62-71, Jan 1998.

- [143] Watkins, C. J. C. H., "Learning from Delayed Rewards," Ph.D. Thesis, Cambridge University, Cambridge, England, 1989.
- [144] Werbös, P. J., "Approximate Dynamic Programming for Real-Time Control and Neural Modeling," in D. A. White and D. A. Sofge, (eds.), *Handbook of Intelligent Control*, Van Nostrand, NY, 1992.
- [145] D. A. White and D. A. Sofge, (eds.), *Handbook of Intelligent Control*, Van Nostrand, NY, 1992.
- [146] P. M. Young. Controller design with mixed uncertainties. In *Proceedings of the American Control Conference*, pages 2333–2337, Baltimore, MD, June 1994.
- [147] P. M. Young, M. P. Newlin, and J. C. Doyle. Practical computation of the mixed μ problem. In *Proceedings of the American Control Conference*, volume 3, pages 2190–2194, Chicago, June 1992.
- [148] P. Young and M.A. Dahleh. "Controller design for mixed ℓ_1 /LMI performance objectives, IEEE Trans. A-C, Oct 1997.
- [149] P. Young and M.A. Dahleh. "Robust ℓ_p stability and performance," *Systems and Control Letters*, Vol 26, pp. 305–312, 1995.
- [150] G. Zames, "On the metric complexity of casual linear systems: ϵ -entropy and ϵ -dimension for continuous-time", *IEEE Trans. on Automatic Control*, Vol. 24, April 1979.
- [151] V. I. Zubov. *Methods of A. M. Lyapunov and their Application*. P. Noordhoof, Ltd., Groningen, The Netherlands, 1964.